

10508.204-WO.ST25  
SEQUENCE LISTING

<110> Novozymes A/S  
<120> Protease Variants  
<130> 10508  
<160> 21  
<170> PatentIn version 3.2  
<210> 1  
<211> 1596  
<212> DNA  
<213> Nocardiosis sp. NRRL 18262 ("Protease 10")

<220>  
<221> CDS  
<222> (318)..(1463)

<220>  
<221> sig\_peptide  
<222> (318)..(404)

<220>  
<221> mat\_peptide  
<222> (900)..(1463)

<400> 1  
acgttttggtta cgggtaccgg tgtccgcatg tggccagaat gcccccttgc gacaggggaac 60  
ggattcggtc ggtagcgcac cgactccgac aaccgcgagg tggccgttcg cgtcgccacg 120  
ttctgcgacc gtcacgcgac ccacatcggg gtgacccac cgagctctga atgggtccacc 180  
gttctgacgg tctttccctc accaaaacgt gcacctatgg ttaggacggt gtttaccgaa 240  
tgtctcggtg aacgacaggg gccggacggt attcggcccc gatcccccggt tgatcccccc 300  
aggagagtag ggacccc atg cga ccc tcc ccc gtt gtc tcc gcc atc ggt 350  
Met Arg Pro Ser Pro Val Val Ser Ala Ile Gly  
-190 -185  
  
acg gga gcg ctg gcc ttc ggt ctg gcg ctg tcc ggt acc ccg ggt 395  
Thr Gly Ala Leu Ala Phe Gly Leu Ala Leu Ser Gly Thr Pro Gly  
-180 -175 -170  
  
gcc ctc gcg gcc acc gga gcg ctc ccc cag tca ccc acc ccg gag 440  
Ala Leu Ala Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu  
-165 -160 -155  
  
gcc gac gcg gtc tcc atg cag gag gcg ctc cag cgc gac ctc gac 485  
Ala Asp Ala Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp  
-150 -145 -140  
  
ctg acc tcc gcc gag gcc gag gag ctg ctg gcc gcc cag gac acc 530  
Leu Thr Ser Ala Glu Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr  
-135 -130 -125  
  
gcc ttc gag gtc gac gag gcc gcg gcc gag gcc gcc ggg gac gcc 575  
Ala Phe Glu Val Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala  
-120 -115 -110  
  
tac ggc ggc tcc gtc ttc gac acc gag agc ctg gaa ctg acc gtc ctg 623  
Tyr Gly Gly Ser Val Phe Asp Thr Glu Ser Leu Glu Leu Thr Val Leu  
-105 -100 -95

## 10508.204-WO.ST25

gtc	acc	gat	gcc	gcc	gcg	gtc	gag	gcc	gtg	gag	gcc	acc	ggc	gcc	ggg	671
Val	Thr	Asp	Ala	Ala	Ala	Val	Glu	Ala	Val	Glu	Ala	Thr	Gly	Ala	Gly	
		-90					-85					-80				
acc	gag	ctg	gtc	tcc	tac	ggc	atc	gac	ggg	ctc	gac	gag	atc	gtc	cag	719
Thr	Glu	Leu	Val	Ser	Tyr	Gly	Ile	Asp	Gly	Leu	Asp	Glu	Ile	Val	Gln	
	-75					-70					-65					
gag	ctc	aac	gcc	gcc	gac	gcc	gtt	ccc	ggg	gtg	gtc	ggc	tgg	tac	ccg	767
Glu	Leu	Asn	Ala	Ala	Asp	Ala	Val	Pro	Gly	Val	Val	Gly	Trp	Tyr	Pro	
-60					-55					-50					-45	
gac	gtg	gcg	ggg	gac	acc	gtc	gtc	ctg	gag	gtc	ctg	gag	ggg	tcc	gga	815
Asp	Val	Ala	Gly	Asp	Thr	Val	Val	Leu	Glu	Val	Leu	Glu	Gly	Ser	Gly	
				-40					-35					-30		
gcc	gac	gtc	agc	ggc	ctg	ctc	gcg	gac	gcc	ggc	gtg	gac	gcc	tcg	gcc	863
Ala	Asp	Val	Ser	Gly	Leu	Leu	Ala	Asp	Ala	Gly	Val	Asp	Ala	Ser	Ala	
			-25					-20					-15			
gtc	gag	gtg	acc	acg	agc	gac	cag	ccc	gag	ctc	tac	gcc	gac	atc	atc	911
Val	Glu	Val	Thr	Thr	Ser	Asp	Gln	Pro	Glu	Leu	Tyr	Ala	Asp	Ile	Ile	
	-10						-5				-1	1				
ggg	ggg	ctg	gcc	tac	acc	atg	ggc	ggc	cgc	tgt	tcg	gtc	ggc	ttc	gcg	959
Gly	Gly	Leu	Ala	Tyr	Thr	Met	Gly	Gly	Arg	Cys	Ser	Val	Gly	Phe	Ala	
5					10					15					20	
gcc	acc	aac	gcc	gcc	ggg	cag	ccc	ggg	ttc	gtc	acc	gcc	ggg	cac	tgc	1007
Ala	Thr	Asn	Ala	Ala	Gly	Gln	Pro	Gly	Phe	Val	Thr	Ala	Gly	His	Cys	
				25					30					35		
ggc	cgc	gtg	ggc	acc	cag	gtg	acc	atc	ggc	aac	ggc	agg	ggc	gtc	ttc	1055
Gly	Arg	Val	Gly	Thr	Gln	Val	Thr	Ile	Gly	Asn	Gly	Arg	Gly	Val	Phe	
			40					45					50			
gag	cag	tcc	gtc	ttc	ccc	ggc	aac	gac	gcg	gcc	ttc	gtc	cgc	ggg	acg	1103
Glu	Gln	Ser	Val	Phe	Pro	Gly	Asn	Asp	Ala	Ala	Phe	Val	Arg	Gly	Thr	
		55					60					65				
tcc	aac	ttc	acg	ctg	acc	aac	ctg	gtc	agc	cgc	tac	aac	acc	ggc	ggg	1151
Ser	Asn	Phe	Thr	Leu	Thr	Asn	Leu	Val	Ser	Arg	Tyr	Asn	Thr	Gly	Gly	
	70					75					80					
tac	gcc	acg	gtc	gcc	ggg	cac	aac	cag	gcc	ccc	atc	ggc	tcc	tcc	gtc	1199
Tyr	Ala	Thr	Val	Ala	Gly	His	Asn	Gln	Ala	Pro	Ile	Gly	Ser	Ser	Val	
85					90					95					100	
tgc	cgc	tcc	ggc	tcc	acc	acc	ggg	tgg	cac	tgt	ggc	acc	atc	cag	gcc	1247
Cys	Arg	Ser	Gly	Ser	Thr	Thr	Gly	Trp	His	Cys	Gly	Thr	Ile	Gln	Ala	
				105					110					115		
cgc	ggc	cag	tcg	gtg	agc	tac	ccc	gag	ggc	acc	gtc	acc	aac	atg	acc	1295
Arg	Gly	Gln	Ser	Val	Ser	Tyr	Pro	Glu	Gly	Thr	Val	Thr	Asn	Met	Thr	
			120					125					130			
cgg	acc	acc	gtg	tgc	gcc	gag	ccc	ggc	gac	tcc	ggc	ggc	tcc	tac	atc	1343
Arg	Thr	Thr	Val	Cys	Ala	Glu	Pro	Gly	Asp	Ser	Gly	Gly	Ser	Tyr	Ile	
		135					140					145				
tcc	ggc	acc	cag	gcc	cag	ggc	gtg	acc	tcc	ggc	ggc	tcc	ggc	aac	tgc	1391
Ser	Gly	Thr	Gln	Ala	Gln	Gly	Val	Thr	Ser	Gly	Gly	Ser	Gly	Asn	Cys	
	150					155					160					
cgc	acc	ggc	ggg	acc	acc	ttc	tac	cag	gag	gtc	acc	ccc	atg	gtg	aac	1439
Arg	Thr	Gly	Gly	Thr	Thr	Phe	Tyr	Gln	Glu	Val	Thr	Pro	Met	Val	Asn	
165					170					175					180	

## 10508.204-WO.ST25

tcc tgg ggc gtc cgt ctc cgg acc tgatccccgc gggtccaggc ggaccgacgg 1493  
 Ser Trp Gly Val Arg Leu Arg Thr  
 185

tcgtgacctg agtaccaggc gtccccgccg cttccagcgg cgtccgcacc ggggtgggac 1553  
 cgggcgtggc cacggcccca cccgtgaccg gaccgcccgg cta 1596

<210> 2  
 <211> 382  
 <212> PRT  
 <213> Nocardiosis sp. NRRL 18262 ("Protease 10")

<400> 2

Met Arg Pro Ser Pro Val Val Ser Ala Ile Gly Thr Gly Ala Leu  
 -190 -185 -180

Ala Phe Gly Leu Ala Leu Ser Gly Thr Pro Gly Ala Leu Ala Ala  
 -175 -170 -165

Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val  
 -160 -155 -150

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala  
 -145 -140 -135

Glu Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val  
 -130 -125 -120

Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser  
 -115 -110 -105

Val Phe Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ala  
 -100 -95 -90

Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu Val  
 -85 -80 -75

Ser Tyr Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn Ala  
 -70 -65 -60

Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly  
 -55 -50 -45

Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Ser  
 -40 -35 -30 -25

Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr  
 -20 -15 -10

Thr Ser Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu Ala  
 -5 -1 1 5

10508.204-WO.ST25

Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn Ala  
10 15 20  
Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val Gly  
25 30 35 40  
Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser Val  
45 50 55  
Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe Thr  
60 65 70  
Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr Val  
75 80 85  
Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser Gly  
90 95 100  
Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln Ser  
105 110 115 120  
Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr Val  
125 130 135  
Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr Gln  
140 145 150  
Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly Gly  
155 160 165  
Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly Val  
170 175 180  
Arg Leu Arg Thr  
185

<210> 3  
<211> 1065  
<212> DNA  
<213> Nocardiosis dassonvillei subspecies dassonvillei DSM 43235 ("Protease 18")

<220>  
<221> CDS  
<222> (1)..(1062)

<220>  
<221> mat\_peptide  
<222> (499)..(1062)

<400> 3  
gct ccg gcc ccc gtc ccc cag acc ccc gtc gcc gac gac agc gcc 45  
Ala Pro Ala Pro Val Pro Gln Thr Pro Val Ala Asp Asp Ser Ala  
-165 -160 -155

## 10508.204-WO.ST25

gcc agc atg acc gag gcg ctc aag cgc gac ctc gac ctc acc tcg Ala Ser Met Thr Glu Ala Leu Lys Arg Asp Leu Asp Leu Thr Ser -150 -145 -140	90
gcc gag gcc gag gag ctt ctc tcg gcg cag gaa gcc gcc atc gag Ala Glu Ala Glu Glu Leu Leu Ser Ala Gln Glu Ala Ala Ile Glu -135 -130 -125	135
acc gac gcc gag gcc acc gag gcc gcg ggc gag gcc tac ggc ggc Thr Asp Ala Glu Ala Thr Glu Ala Ala Gly Glu Ala Tyr Gly Gly -120 -115 -110	180
tca ctg ttc gac acc gag acc ctc gaa ctc acc gtg ctg gtc acc gac Ser Leu Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp -105 -100 -95	228
gcc tcc gcc gtc gag gcg gtc gag gcc acc gga gcc cag gcc acc gtc Ala Ser Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gln Ala Thr Val -90 -85 -80 -75	276
gtc tcc cac ggc acc gag ggc ctg acc gag gtc gtg gag gac ctc aac Val Ser His Gly Thr Glu Gly Leu Thr Glu Val Val Glu Asp Leu Asn -70 -65 -60	324
ggc gcc gag gtt ccc gag agc gtc ctc ggc tgg tac ccg gac gtg gag Gly Ala Glu Val Pro Glu Ser Val Leu Gly Trp Tyr Pro Asp Val Glu -55 -50 -45	372
agc gac acc gtc gtg gtc gag gtg ctg gag ggc tcc gac gcc gac gtc Ser Asp Thr Val Val Val Glu Val Leu Glu Gly Ser Asp Ala Asp Val -40 -35 -30	420
gcc gcc ctg ctc gcc gac gcc ggt gtg gac tcc tcc tcg gtc cgg gtg Ala Ala Leu Leu Ala Asp Ala Gly Val Asp Ser Ser Ser Val Arg Val -25 -20 -15	468
gag gag gcc gag gag gcc ccg cag gtc tac gcc gac atc atc ggc ggc Glu Glu Ala Glu Glu Ala Pro Gln Val Tyr Ala Asp Ile Ile Gly Gly -10 -5 -1 1 5	516
ctg gcc tac tac atg ggc ggc cgc tgc tcc gtc ggc ttc gcc gcg acc Leu Ala Tyr Tyr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr 10 15 20	564
aac agc gcc ggt cag ccc ggt ttc gtc acc gcc ggc cac tgc ggc acc Asn Ser Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr 25 30 35	612
gtc ggc acc ggc gtg acc atc ggc aac ggc acc ggc acc ttc cag aac Val Gly Thr Gly Val Thr Ile Gly Asn Gly Thr Gly Thr Phe Gln Asn 40 45 50	660
tcg gtc ttc ccc ggc aac gac gcc gcc ttc gtc cgc ggc acc tcc aac Ser Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn 55 60 65 70	708
ttc acc ctg acc aac ctg gtc tcg cgc tac aac tcc ggc ggc tac cag Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Gln 75 80 85	756
tcg gtg acc ggt acc agc cag gcc ccg gcc ggc tcg gcc gtg tgc cgc Ser Val Thr Gly Thr Ser Gln Ala Pro Ala Gly Ser Ala Val Cys Arg 90 95 100	804
tcc ggc tcc acc acc ggc tgg cac tgc ggc acc atc cag gcc cgc aac Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn 105 110 115	852

## 10508.204-WO.ST25

cag acc gtg cgc tac ccg cag ggc acc gtc tac tcg ctc acc cgc acc 900  
 Gln Thr Val Arg Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr  
 120 125 130  
 aac gtg tgc gcc gag ccc ggc gac tcc ggc ggt tcg ttc atc tcc ggc 948  
 Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser Gly  
 135 140 145 150  
 tcg cag gcc cag ggc gtc acc tcc ggc ggc tcc ggc aac tgc tcc gtc 996  
 Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Val  
 155 160 165  
 ggc ggc acg acc tac tac cag gag gtc acc ccg atg atc aac tcc tgg 1044  
 Gly Gly Thr Thr Tyr Tyr Gln Glu Val Thr Pro Met Ile Asn Ser Trp  
 170 175 180  
 ggt gtc agg atc cgg acc taa 1065  
 Gly Val Arg Ile Arg Thr  
 185

<210> 4  
 <211> 354  
 <212> PRT  
 <213> Nocardiosis dassonvillei subspecies dassonvillei DSM 43235 ("Protease 18")

<400> 4

Ala Pro Ala Pro Val Pro Gln Thr Pro Val Ala Asp Asp Ser Ala  
 -165 -160 -155  
 Ala Ser Met Thr Glu Ala Leu Lys Arg Asp Leu Asp Leu Thr Ser  
 -150 -145 -140  
 Ala Glu Ala Glu Glu Leu Leu ser Ala Gln Glu Ala Ala Ile Glu  
 -135 -130 -125  
 Thr Asp Ala Glu Ala Thr Glu Ala Ala Gly Glu Ala Tyr Gly Gly  
 -120 -115 -110  
 Ser Leu Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp  
 -105 -100 -95  
 Ala Ser Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gln Ala Thr Val  
 -90 -85 -80 -75  
 Val Ser His Gly Thr Glu Gly Leu Thr Glu Val Val Glu Asp Leu Asn  
 -70 -65 -60  
 Gly Ala Glu Val Pro Glu Ser Val Leu Gly Trp Tyr Pro Asp Val Glu  
 -55 -50 -45  
 Ser Asp Thr Val Val Val Glu Val Leu Glu Gly Ser Asp Ala Asp Val  
 -40 -35 -30  
 Ala Ala Leu Leu Ala Asp Ala Gly Val Asp Ser Ser Ser Val Arg Val  
 -25 -20 -15

10508.204-WO.ST25

Glu Glu Ala Glu Glu Ala Pro Gln Val Tyr Ala Asp Ile Ile Gly Gly  
 -10 -5 -1 1 5  
 Leu Ala Tyr Tyr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr  
 10 15 20  
 Asn Ser Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr  
 25 30 35  
 Val Gly Thr Gly Val Thr Ile Gly Asn Gly Thr Gly Thr Phe Gln Asn  
 40 45 50  
 Ser Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn  
 55 60 65 70  
 Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Gln  
 75 80 85  
 Ser Val Thr Gly Thr Ser Gln Ala Pro Ala Gly Ser Ala Val Cys Arg  
 90 95 100  
 Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn  
 105 110 115  
 Gln Thr Val Arg Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr  
 120 125 130  
 Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser Gly  
 135 140 145 150  
 Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Val  
 155 160 165  
 Gly Gly Thr Thr Tyr Tyr Gln Glu Val Thr Pro Met Ile Asn Ser Trp  
 170 175 180  
 Gly Val Arg Ile Arg Thr  
 185

<210> 5  
 <211> 1062  
 <212> DNA  
 <213> Nocardiosis prasina DSM 15648 ("Protease 11")

<220>  
 <221> CDS  
 <222> (1)..(1059)  
 <220>  
 <221> mat\_peptide  
 <222> (496)..(1059)



## 10508.204-WO.ST25

<400> 5  
 gcc acc gga ccg ctc ccc cag tca ccc acc ccg gag gcc gac gcc 45  
 Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala  
 -165 -160 -155

gtc tcc atg cag gag gcg ctc cag cgc gac ctc ggc ctg acc ccg 90  
 Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro  
 -150 -145 -140

ctt gag gcc gat gaa ctg ctg gcc gcc cag gac acc gcc ttc gag 135  
 Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu  
 -135 -130 -125

gtc gac gag gcc gcg gcc gcg gcc gcc ggg gac gcc tac ggc ggc 180  
 Val Asp Glu Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly  
 -120 -115 -110

tcc gtc ttc gac acc gag acc ctg gaa ctg acc gtc ctg gtc acc gac 228  
 Ser Val Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp  
 -105 -100 -95

gcc gcc tcg gtc gag gct gtg gag gcc acc ggc gcg ggt acc gaa ctc 276  
 Ala Ala Ser Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu  
 -85 -80 -75

gtc tcc tac ggc atc gag ggc ctc gac gag atc atc cag gat ctc aac 324  
 Val Ser Tyr Gly Ile Glu Gly Leu Asp Glu Ile Ile Gln Asp Leu Asn  
 -70 -65 -60

gcc gcc gac gcc gtc ccc ggc gtg gtc ggc tgg tac ccg gac gtg gcg 372  
 Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala  
 -55 -50 -45

ggt gac acc gtc gtc ctg gag gtc ctg gag ggt tcc gga gcc gac gtg 420  
 Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val  
 -40 -35 -30

agc ggc ctg ctc gcc gac gcc ggc gtg gac gcc tcg gcc gtc gag gtg 468  
 Ser Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val  
 -25 -20 -15 -10

acc agc agt gcg cag ccc gag ctc tac gcc gac atc atc ggc ggt ctg 516  
 Thr Ser Ser Ala Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu  
 -5 -1 1 5

gcc tac acc atg ggc ggc cgc tgt tcg gtc gga ttc gcg gcc acc aac 564  
 Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn  
 10 15 20

gcc gcc ggt cag ccc gga ttc gtc acc gcc ggt cac tgt ggc cgc gtg 612  
 Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val  
 25 30 35

ggc acc cag gtg agc atc ggc aac ggc cag ggc gtc ttc gag cag tcc 660  
 Gly Thr Gln Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu Gln Ser  
 40 45 50 55

atc ttc ccg ggc aac gac gcc gcc ttc gtc cgc ggc acg tcc aac ttc 708  
 Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe  
 60 65 70

acg ctg acc aac ctg gtc agc cgc tac aac acc ggc ggt tac gcc acc 756  
 Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr  
 75 80 85

gtc gcc ggc cac aac cag gcg ccc atc ggc tcc tcc gtc tgc cgc tcc 804  
 Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser  
 90 95 100



## 10508.204-WO.ST25

ggc tcc acc acc ggc tgg cac tgc ggc acc atc cag gcc cgc ggc cag Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln 105 110 115	852
tcg gtg agc tac ccc gag ggc acc gtc acc aac atg acc cgg acc acc Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr 120 125 130 135	900
gtg tgc gcc gag ccc ggc gac tcc ggc ggc tcc tac atc tcc ggc aac Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Asn 140 145 150	948
cag gcc cag ggc gtc acc tcc ggc ggc tcc ggc aac tgc cgc acc ggc Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly 155 160 165	996
ggg acc acc ttc tac cag gag gtc acc ccc atg gtg aac tcc tgg ggc Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly 170 175 180	1044
gtc cgt ctc cgg acc taa Val Arg Leu Arg Thr 185	1062

<210> 6  
 <211> 353  
 <212> PRT  
 <213> Nocardiosis prasina DSM 15648 ("Protease 11")

<400> 6

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala -165 -160 -155
Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro -150 -145 -140
Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu -135 -130 -125
Val Asp Glu Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly -120 -115 -110
Ser Val Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp -105 -100 -95 -90
Ala Ala Ser Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu -85 -80 -75
Val Ser Tyr Gly Ile Glu Gly Leu Asp Glu Ile Ile Gln Asp Leu Asn -70 -65 -60
Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala -55 -50 -45
Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val -40 -35 -30

## 10508.204-WO.ST25

Ser Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val  
-25 -20 -15 -10

Thr Ser Ser Ala Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu  
-5 -1 1 5

Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn  
10 15 20

Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val  
25 30 35

Gly Thr Gln Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu Gln Ser  
40 45 50 55

Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe  
60 65 70

Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr  
75 80 85

Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser  
90 95 100

Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln  
105 110 115

Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr  
120 125 130 135

Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Asn  
140 145 150

Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly  
155 160 165

Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly  
170 175 180

Val Arg Leu Arg Thr  
185

<210> 7  
<211> 1062  
<212> DNA  
<213> Nocardiosis prasina DSM 15649 ("Protease 35")

<220>  
<221> CDS  
<222> (1)..(1059)

<220>

## 10508.204-WO.ST25

&lt;221&gt; mat\_peptide

&lt;222&gt; (496)..(1059)

&lt;400&gt; 7

gcc acc gga cca ctc ccc cag tca ccc acc ccg gag gcc gac gcc Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala -165 -160 -155	45
gtc tcc atg cag gag gcg ctc cag cgc gac ctc ggc ctg acc ccg Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro -150 -145 -140	90
ctt gag gcc gat gaa ctg ctg gcc gcc cag gac acc gcc ttc gag Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu -135 -130 -125	135
gtc gac gag gcc gcg gcc gag gcc gcc ggt gac gcc tac ggc ggc Val Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly -120 -115 -110	180
tcc gtc ttc gac acc gag acc ctg gaa ctg acc gtc ctg gtc acc gac Ser Val Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp -105 -100 -95 -90	228
tcc gcc gcg gtc gag gcg gtg gag gcc acc ggc gcc ggg acc gaa ctg Ser Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu -85 -80 -75	276
gtc tcc tac ggc atc acg ggc ctc gac gag atc gtc gag gag ctc aac Val Ser Tyr Gly Ile Thr Gly Leu Asp Glu Ile Val Glu Glu Leu Asn -70 -65 -60	324
gcc gcc gac gcc gtt ccc ggc gtg gtc ggc tgg tac ccg gac gtc gcg Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala -55 -50 -45	372
ggt gac acc gtc gtg ctg gag gtc ctg gag ggt tcc ggc gcc gac gtg Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val -40 -35 -30	420
ggc ggc ctg ctc gcc gac gcc ggc gtg gac gcc tcg gcg gtc gag gtg Gly Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val -25 -20 -15 -10	468
acc acc acc gag cag ccc gag ctg tac gcc gac atc atc ggc ggt ctg Thr Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu -5 -1 1 5	516
gcc tac acc atg ggc ggc cgc tgt tcg gtc ggc ttc gcg gcc acc aac Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn 10 15 20	564
gcc gcc ggt cag ccc ggc ttc gtc acc gcc ggt cac tgt ggc cgc gtg Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val 25 30 35	612
ggc acc cag gtg acc atc ggc aac ggc cgg ggc gtc ttc gag cag tcc Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser 40 45 50 55	660
atc ttc ccg ggc aac gac gcc gcc ttc gtc cgc gga acg tcc aac ttc Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe 60 65 70	708
acg ctg acc aac ctg gtc agc cgc tac aac acc ggc ggc tac gcc acc Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr 75 80 85	756

10508.204-WO.ST25

gtc	gcc	ggt	cac	aac	cag	gcg	ccc	atc	ggc	tcc	tcc	gtc	tgc	cgc	tcc	804
Val	Ala	Gly	His	Asn	Gln	Ala	Pro	Ile	Gly	Ser	Ser	Val	Cys	Arg	Ser	
		90					95					100				
ggc	tcc	acc	acc	ggt	tgg	cac	tgc	ggc	acc	atc	cag	gcc	cgc	ggc	cag	852
Gly	Ser	Thr	Thr	Gly	Trp	His	Cys	Gly	Thr	Ile	Gln	Ala	Arg	Gly	Gln	
	105					110					115					
tcg	gtg	agc	tac	ccc	gag	ggc	acc	gtc	acc	aac	atg	acg	cgg	acc	acc	900
Ser	Val	Ser	Tyr	Pro	Glu	Gly	Thr	Val	Thr	Asn	Met	Thr	Arg	Thr	Thr	
	120				125					130					135	
gtg	tgc	gcc	gag	ccc	ggc	gac	tcc	ggc	ggc	tcc	tac	atc	tcc	ggc	aac	948
Val	Cys	Ala	Glu	Pro	Gly	Asp	Ser	Gly	Gly	Ser	Tyr	Ile	Ser	Gly	Asn	
				140				145						150		
cag	gcc	cag	ggc	gtc	acc	tcc	ggc	ggc	tcc	ggc	aac	tgc	cgc	acc	ggc	996
Gln	Ala	Gln	Gly	Val	Thr	Ser	Gly	Gly	Ser	Gly	Asn	Cys	Arg	Thr	Gly	
			155					160					165			
ggg	acc	acc	ttc	tac	cag	gag	gtc	acc	ccc	atg	gtg	aac	tcc	tgg	ggc	1044
Gly	Thr	Thr	Phe	Tyr	Gln	Glu	Val	Thr	Pro	Met	Val	Asn	Ser	Trp	Gly	
		170					175					180				
gtc	cgt	ctc	cgg	acc	taa											1062
Val	Arg	Leu	Arg	Thr												
	185															

<210> 8  
 <211> 353  
 <212> PRT  
 <213> Nocardiosis prasina DSM 15649 ("Protease 35")

<400> 8

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala  
 -165 -160 -155

Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro  
 -150 -145 -140

Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu  
 -135 -130 -125

Val Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly  
 -120 -115 -110

Ser Val Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp  
 -105 -100 -95 -90

Ser Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu  
 -85 -80 -75

Val Ser Tyr Gly Ile Thr Gly Leu Asp Glu Ile Val Glu Glu Leu Asn  
 -70 -65 -60

Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala  
 -55 -50 -45

10508.204-WO.ST25

Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val  
 -40 -35 -30

Gly Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val  
 -25 -20 -15 -10

Thr Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu  
 -5 -1 1 5

Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn  
 10 15 20

Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val  
 25 30 35

Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser  
 40 45 50 55

Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe  
 60 65 70

Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr  
 75 80 85

Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser  
 90 95 100

Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln  
 105 110 115

Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr  
 120 125 130 135

Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Asn  
 140 145 150

Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly  
 155 160 165

Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly  
 170 175 180

Val Arg Leu Arg Thr  
 185

<210> 9  
 <211> 1068  
 <212> DNA  
 <213> Nocardiopsis alba DSM 15647 ("Protease 08")

<220>  
 <221> CDS

10508.204-WO.ST25

&lt;222&gt; (1)..(1065)

&lt;220&gt;

&lt;221&gt; mat\_peptide

&lt;222&gt; (502)..(1065)

&lt;400&gt; 9

gcg acc ggc ccc ctc ccc cag tcc ccc acc ccg gat gaa gcc gag	45
Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Asp Glu Ala Glu	
-165 -160 -155	
gcc acc acc atg gtc gag gcc ctc cag cgc gac ctc ggc ctg tcc	90
Ala Thr Thr Met Val Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser	
-150 -145 -140	
ccc tct cag gcc gac gag ctc ctc gag gcg cag gcc gag tcc ttc	135
Pro Ser Gln Ala Asp Glu Leu Leu Glu Ala Gln Ala Glu Ser Phe	
-135 -130 -125	
gag atc gac gag gcc gcc acc gcg gcc gca gcc gac tcc tac ggc	180
Glu Ile Asp Glu Ala Ala Thr Ala Ala Ala Asp Ser Tyr Gly	
-120 -115 -110	
ggc tcc atc ttc gac acc gac agc ctc acc ctg acc gtc ctg gtc acc	228
Gly Ser Ile Phe Asp Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr	
-105 -100 -95	
gac gcc tcc gcc gtc gag gcg gtc gag gcc gcc ggc gcc gag gcc aag	276
Asp Ala Ser Ala Val Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys	
-90 -85 -80	
gtg gtc tcg cac ggc atg gag ggc ctg gag gag atc gtc gcc gac ctg	324
Val Val Ser His Gly Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu	
-75 -70 -65 -60	
aac gcg gcc gac gct cag ccc ggc gtc gtg ggc tgg tac ccc gac atc	372
Asn Ala Ala Asp Ala Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile	
-55 -50 -45	
cac tcc gac acg gtc gtc ctc gag gtc ctc gag ggc tcc ggt gcc gac	420
His Ser Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp	
-40 -35 -30	
gtg gac tcc ctg ctc gcc gac gcc ggt gtg gac acc gcc gac gtc aag	468
Val Asp Ser Leu Leu Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys	
-25 -20 -15	
gtg gag agc acc acc gag cag ccc gag ctg tac gcc gac atc atc ggc	516
Val Glu Ser Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly	
-10 -5 -1 1 5	
ggt ctc gcc tac acc atg ggt ggg cgc tgc tcg gtc ggc ttc gcg gcc	564
Gly Leu Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala	
10 15 20	
acc aac gcc tcc ggc cag ccc ggg ttc gtc acc gcc ggc cac tgc ggc	612
Thr Asn Ala Ser Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly	
25 30 35	
acc gtc ggc acc ccg gtc agc atc ggc aac ggc cag ggc gtc ttc gag	660
Thr Val Gly Thr Pro Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu	
40 45 50	
cgt tcc gtc ttc ccc ggc aac gac tcc gcc ttc gtc cgc ggc acc tcg	708
Arg Ser Val Phe Pro Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser	
55 60 65	
aac ttc acc ctg acc aac ctg gtc agc cgc tac aac acc ggt ggt tac	756

10508.204-WO.ST25

Asn 70	Phe	Thr	Leu	Thr	Asn 75	Leu	Val	Ser	Arg	Tyr 80	Asn	Thr	Gly	Gly	Tyr 85	
gcg Ala	acc Thr	gtc Val	tcc Ser	ggc Gly 90	tcc Ser	tcg Ser	cag Gln	gcg Ala	gcg Ala 95	atc Ile	ggc Gly	tcg Ser	cag Gln	atc Ile 100	tgc Cys	804
cgt Arg	tcc Ser	ggc Gly	tcc Ser 105	acc Thr	acc Thr	ggc Gly	tgg Trp	cac His 110	tgc Cys	ggc Gly	acc Thr	gtc Val	cag Gln 115	gcc Ala	cgc Arg	852
ggc Gly	cag Gln	acg Thr 120	gtg Val	agc Ser	tac Tyr	ccc Pro	cag Gln 125	ggc Gly	acc Thr	gtg Val	cag Gln	aac Asn 130	ctg Leu	acc Thr	cgc Arg	900
acc Thr	aac Asn 135	gtc Val	tgc Cys	gcc Ala	gag Glu	ccc Pro 140	ggc Gly	gac Asp	tcc Ser	ggc Gly	ggc Gly 145	tcc Ser	ttc Phe	atc Ile	tcc Ser	948
ggc Gly 150	agc Ser	cag Gln	gcc Ala	cag Gln	ggc Gly 155	gtc Val	acc Thr	tcc Ser	ggc Gly	ggc Gly 160	tcc Ser	ggc Gly	aac Asn	tgc Cys	tcc Ser 165	996
ttc Phe	ggc Gly	ggc Gly	acc Thr	acc Thr 170	tac Tyr	tac Tyr	cag Gln	gag Glu	gtc Val 175	aac Asn	ccg Pro	atg Met	ctg Leu	agc Ser 180	agc Ser	1044
tgg Trp	ggc Gly	ctg Leu	acc Thr 185	ctg Leu	cgc Arg	acc Thr	tga									1068

<210> 10  
 <211> 355  
 <212> PRT  
 <213> Nocardiosis alba DSM 15647 ("Protease 08")

<400> 10

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Asp Glu Ala Glu  
 -165 -160 -155

Ala Thr Thr Met Val Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser  
 -150 -145 -140

Pro Ser Gln Ala Asp Glu Leu Leu Glu Ala Gln Ala Glu Ser Phe  
 -135 -130 -125

Glu Ile Asp Glu Ala Ala Thr Ala Ala Ala Ala Asp Ser Tyr Gly  
 -120 -115 -110

Gly Ser Ile Phe Asp Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr  
 -105 -100 -95

Asp Ala Ser Ala Val Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys  
 -90 -85 -80

Val Val Ser His Gly Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu  
 -75 -70 -65 -60

Asn Ala Ala Asp Ala Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile



10508.204-WO.ST25

-55

-50

-45

His Ser Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp  
 -40 -35 -30

Val Asp Ser Leu Leu Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys  
 -25 -20 -15

Val Glu Ser Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly  
 -10 -5 -1 1 5

Gly Leu Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala  
 10 15 20

Thr Asn Ala Ser Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly  
 25 30 35

Thr Val Gly Thr Pro Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu  
 40 45 50

Arg Ser Val Phe Pro Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser  
 55 60 65

Asn Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr  
 70 75 80 85

Ala Thr Val Ser Gly Ser Ser Gln Ala Ala Ile Gly Ser Gln Ile Cys  
 90 95 100

Arg Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Val Gln Ala Arg  
 105 110 115

Gly Gln Thr Val Ser Tyr Pro Gln Gly Thr Val Gln Asn Leu Thr Arg  
 120 125 130

Thr Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser  
 135 140 145

Gly Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser  
 150 155 160 165

Phe Gly Gly Thr Thr Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser  
 170 175 180

Trp Gly Leu Thr Leu Arg Thr  
 185

<210> 11  
 <211> 27  
 <212> DNA  
 <213> Artificial

10508.204-WO.ST25

<220>  
 <223> Primer

<400> 11  
 ccgattatgg agcggattga acatgcg

27

<210> 12  
 <211> 31  
 <212> DNA  
 <213> Artificial

<220>  
 <223> Primer

<400> 12  
 gtgaccatcg gcgacggcag gggcgtcttc g

31

<210> 13  
 <211> 10172  
 <212> DNA  
 <213> Artificial

<220>  
 <223> Expression Cassette

<220>  
 <221> source  
 <222> (1)..(3323)  
 <223> Bacillus subtilis genome sequence including  
 yfmH-yfmD-yfmC-yfmB-yfmA genes

<220>  
 <221> misc\_recomb  
 <222> (3561)..(4208)  
 <223> Cat gene providing chloramphenicol resistance

<220>  
 <221> promoter  
 <222> (4523)..(5633)  
 <223> Triple PamyL-sCBAN-CryIIIA promoter including mRNA stabilizing  
 sequence

<220>  
 <221> sig\_peptide  
 <222> (5658)..(5738)

<220>  
 <221> CDS  
 <222> (5658)..(6797)

<220>  
 <221> mat\_peptide  
 <222> (6234)..(6797)

<220>  
 <221> source  
 <222> (6839)..(7540)  
 <223> Part of Bacillus subtilis pectate lyase gene

<220>  
 <221> source  
 <222> (7541)..(10172)  
 <223> Bacillus subtilis genome DNA including yflS-citM genes

<400> 13

10508.204-WO.ST25

gagtatcgcc agtaaggggc gtttttgttt tctggttggt ttcttcattt caggtttcgc	60
cctttccttg ccaaataata gaaaaacggc gttccgataa tcgcggtgac aatgccgacc	120
ggtgattcat aaggaaatgc aatccatctg gccagaacat ctgcgtacac cagcaaaatg	180
gcaccgaaca gtgccgaaaa cggaagcacg tattgataat gttctccgat cagcttgcg	240
acaatatgcg ggacgagcag cccgacaaag ccaatcggcc cggcgacggc tacggaagcg	300
ccggaaagaa ttaaaataat caaactgatc agaatcctga tgccgttcat attttgtcca	360
agcccttttg ctgtttcgtc tccgagaccg agaacagaaa cagaaccgga aaatacgagg	420
gcaagcccga tgccaatgac agaaaaagga gcgatgggta tgacgtcctg ccagttgctg	480
ccgtcgattg cgctgtcat ccagtacaga acatcctcac ctgactcatt taaaataatg	540
atggcctgtg tcatagagga gaggaacaag tgcacggcca ttcctgacag cgccagcttg	600
acaggcgtca ttccgccgga tgaggcaatc atatacacia tcgcgccgcc tgctgccgca	660
cccgcaaaag cgaatataac agatgaatag ggcgatgccg gcagaatgac gagagaagca	720
acaacaaaaa gcgatgcacc cgcattcaca ccgaaaattt ggggtgaagc cagaggattt	780
ctggtcatag cctgcatcag cgcccctgct acagctaggc tggcgccgac aaaaacgccg	840
attaatgtgc ggggaaggcg aagagtagag atgatgagct gttcctttga accgtcccat	900
acaaaaagat atttcaatga atctatgatg ctgatgtctg aggctcctac tgaaagattc	960
agccaagcc caaatataaa aataatcagt gcaatgataa acatcatcag tcttgatgat	1020
gagcgccgtt tggctgaatg atacaacagt ctacttcct tactgctct ggttgcaaaa	1080
acgaagaagc aaggattccc ctgccttctc atttgtccta tttattatac acttttttaa	1140
gcacatcttt ggcgcttggt tcactagact tgatgcctct gaatcttgtc caagtgtcac	1200
ggtccgcatc atagacttgt ccatttttca ccgctttgag atttttccag agcgggttcg	1260
ttttccactc atctacaatg gttttgcctt cggttgctga gatgaacaaa atatcaggat	1320
cgattttgct caattgctca aggctgacct cttgataggc gttatctgac ttcacagcgt	1380
gtgtaaagcc tagcatttta aagattttctc cgtcatagga tgatgatgta tgaagctgga	1440
aggaatccgc tcttgcaacg ccgagaacga tggtgcgggt ttcactcttc ggaagttcgg	1500
cttttagatc gttgatgact tttttgtgct cggcaagctt ttcttttcct tcatcttctt	1560
tatttaatgc tttagcaatg gtcgtaaagc tgtcgatcgt ttcgtcatat gtcgcttcac	1620
ggctttttaa ttcaatcgtc ggggcgattt ttttcagctg ttataaaatg tttttatggc	1680
gctcagcgtc agcgatgatt aaatcaggct tcaaggaact gatgacctca agattgggtt	1740
cgctgcgtgt gcctacagat gtgtaatcaa tggagctgcc gacaagcttt ttaatcatat	1800
cttttttggt gtcactctgc atgcccaccg gcgtaatgcc gagattgtga acggcatcca	1860
agaatgaaag ctcaagcaca accacccgct taggtgtgcc gcttactgtc gtttttcctt	1920
cttcgtcatg gatcactctg gaatccttag actcgctttt gccgcttccg ttgttattct	1980
ggcttgatga acagccggat acaatgaggc aggcgagcaa taaaacactc atgatggcaa	2040

## 10508.204-WO.ST25

tcaacttggtt	agaatagggtg	cgcatgtcat	tcttcctttt	ttcagattta	gtaatgagaa	2100
tcattatcac	atgtaacact	ataatagcat	ggcttatcat	gtcaatattt	ttttagtaaa	2160
gaaagctgcg	tttttactgc	tttctcatga	aagcatcatc	agacacaaat	aagtggatatg	2220
cagcgttacc	gtgtcttcga	gacaaaaacg	catgggcggt	ggcttttagag	gtttcgaaca	2280
tatcagcagt	gacataagga	aggagagtgc	tgagataacc	ggacaatttc	ttttctattt	2340
catctgttag	tgcaaattca	atgtcgccga	tattcatgat	aatcgagaaa	acaaagtcga	2400
tatcgatatg	aaaatgttcc	tcggcaaaaa	ccgcaagctc	gtgaattcct	ggtgaacatc	2460
cggcacgctt	atggaaaatc	tgtttgacta	aatcactcac	aatccaagca	ttgtattgct	2520
gttctgggtga	aaagtattgc	attagacata	cctcctgctc	gtacggataa	aggcagcggt	2580
tcatggtcgt	gtgctccgtg	cagcggcttc	tccttaattt	tgatttttct	gaaaatagggt	2640
cccgttccta	tcactttacc	atggacggaa	aacaaatagc	tactaccatt	cctcctgttt	2700
ttctcttcaa	tgttctggaa	tctgtttcag	gtacagacga	tcgggtatga	aagaaatata	2760
gaaaacatga	aggaggaata	tcgacatgaa	accagttgta	aaagagtata	caaatgacga	2820
acagctcatg	aaagatgtag	aggaattgca	gaaaatgggt	gttgcgaaag	aggatgtata	2880
cgtcttagct	cacgacgatg	acagaacgga	acgcctggct	gacaacacga	acgccaacac	2940
gatcggagcc	aaagaaacag	gtttcaagca	cgcggtggga	aatatcttca	ataaaaaagg	3000
agacgagctc	cgcaataaaa	ttcacgaaat	cggtttttct	gaagatgaag	ccgctcaatt	3060
tgaaaaacgc	ttagatgaag	gaaaagtgct	tctctttgtg	acagataacg	aaaaagtgaa	3120
agcttgggca	taaagcaagg	aaaaaaccaa	aaggccaatg	tcggcctttt	ggtttttttg	3180
cggctcttgc	ggtgggattt	tgcagaatgc	cgcaatagga	tagcggaaaca	ttttcggttc	3240
tgaatgtccc	tcaatttgct	attatatttt	tgtgataaat	tggaataaaa	tctcacaaaa	3300
tagaaaatgg	gggtacatag	tggccatcat	ggccagctag	catgcacatg	ggatctggga	3360
ccaataataa	tgactagaga	agaaagaatg	aagattgttc	atgaaattaa	ggaacgaata	3420
ttggataaag	tggggatatt	ttaaaatata	tatttatgtt	acagtaatat	tgacttttaa	3480
aaaaggattg	attctaagaa	gaaagcagac	aagtaagcct	cctaaattca	ctttagataa	3540
aaatttagga	ggcatatcaa	atgaacttta	ataaaattga	tttagacaat	tggaagagaa	3600
aagagatatt	taatcattat	ttgaaccaac	aaacgacttt	tagtataacc	acagaaattg	3660
atattagtgt	tttataccga	aacataaaaac	aagaaggata	taaattttac	cctgcattta	3720
ttttcttagt	gacaaggggtg	ataaactcaa	atacagcttt	tagaactggt	tacaatagcg	3780
acggagagtt	aggttattgg	gataagttag	agccacttta	tacaattttt	gatggtgtat	3840
ctaaaacatt	ctctgggtatt	tggactcctg	taaagaatga	cttcaaagag	ttttatgatt	3900
tatacctttc	tgatgtagag	aaatataatg	gttcgggggaa	attgtttccc	aaaacaccta	3960
tacctgaaaa	tgctttttct	ctttctatta	ttccatggac	ttcattttact	gggtttaact	4020
taaatatcaa	taataatagt	aattaccttc	taccatttat	tacagcagga	aaattcatta	4080

## 10508.204-WO.ST25

ataaaggtaa ttcaatatat ttaccgctat ctttacagggt acatcattct gtttgtgatg 4140  
 gttatcatgc aggattgttt atgaactcta ttcaggaatt gtcagatagg cctaatagact 4200  
 ggctttttata atatgagata atgccgactg tacttttttac agtcgggtttt ctaacgatac 4260  
 attaatagggt acgaaaaagc aactttttttt gcgcttaaaa ccagtcatac caataactta 4320  
 agggtaacta gcctcgccgg aaagagcgaa aatgcctcac atttgtgcca cctaaaaagg 4380  
 agcgattttac atatgagtta tgcagtttgt agaatgcaaa aagtgaatc agctggacta 4440  
 aaaggcatgg catgccttcg atagtattatt aatattagtg gagctcagtg agagcgaagc 4500  
 gaacacttga ttttttaatt ttctatcttt tatagggtcat tagagtatac ttatttgtcc 4560  
 tataaactat ttagcagcat aatagattta ttgaatagggt catttaagtt gagcatatta 4620  
 ggggaggaaa atcttgagaa aatatttgaa gaacccgagg atctagatca ggtaccgcaa 4680  
 cgttcgagaa tgctgctgaa gagattatta aaaagctgaa agcaaaaggc tatcaattgg 4740  
 taactgtatc tcagcttgaa gaagtgaaga agcagagagg ctattgaata aatgagtaga 4800  
 aagcgccata tcggcgcttt tcttttgga gaaaatatag ggaaaatgggt acttggttaa 4860  
 aattcggaat atttatacaa tatcatatgt atcacattga aaggaggggc ctgctgtcca 4920  
 gactgtccgc tgtgtaaaaa aaaggaataa agggggggtg acattatttt actgatattg 4980  
 ataataaat ttgtataaga aaatggaggg gccctcgaaa cgtaagatga aaccttagat 5040  
 aaaagtgctt tttttgttgc aattgaagaa ttattaatgt taagcttaat taaagataat 5100  
 atctttgaat tgtaacgccc ctcaaagta agaactacaa aaaaagaata cgttatatag 5160  
 aaatatgttt gaaccttctt cagattacaa atatattcgg acggactcta cctcaaattgc 5220  
 ttatctaact atagaatgac atacaagcac aaccttgaaa atttgaaaat ataactacca 5280  
 atgaacttgt tcatgtgaat tatcgctgta tttaattttc tcaattcaat atataatatg 5340  
 ccaatacatt gttacaagta gaaattaaga cacccttgat agccttacta tacctaacat 5400  
 gatgtagtat taaatgaata tgtaaataa tttatgataa gaagcgactt atttataatc 5460  
 attacatatt tttctatttg aatgattaag attccaatag aatagtgtat aaattattta 5520  
 tcttgaaagg agggatgcct aaaaacgaag aacattaaaa acatatattt gcaccgtcta 5580  
 atggatttat gaaaaatcat tttatcagtt tgaaaattat gtattatgga gctctgaaaa 5640  
 aaaggagagg ataaaga atg aag aaa ccg ttg ggg aaa att gtc gca agc 5690  
                     Met Lys Lys Pro Leu Gly Lys Ile Val Ala Ser  
                                     -190                                    -185

acc gca cta ctc att tct gtt gct ttt agt tca tcg atc gca tcg 5735  
 Thr Ala Leu Leu Ile Ser Val Ala Phe Ser Ser Ser Ile Ala Ser  
                     -180                    -175                    -170

gct gcc acc gga gcg ctc ccc cag tca ccc acc ccg gag gcc gac 5780  
 Ala Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp  
                     -165                    -160                    -155

gcg gtc tcc atg cag gag gcg ctc cag cgc gac ctc gac ctg acc 5825  
 Ala Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr  
                     -150                    -145                    -140



10508.204-wo.ST25																
tcc	gcc	gag	gcc	gag	gag	ctg	ctg	gcc	gcc	cag	gac	acc	gcc	ttc		5870
Ser	Ala	Glu	Ala	Glu	Glu	Leu	Leu	Ala	Ala	Gln	Asp	Thr	Ala	Phe		
	-135					-130					-125					
gag	gtc	gac	gag	gcc	gcg	gcc	gag	gcc	gcc	ggg	gac	gcc	tac	ggc		5915
Glu	Val	Asp	Glu	Ala	Ala	Ala	Glu	Ala	Ala	Gly	Asp	Ala	Tyr	Gly		
	-120					-115					-110					
ggc	tcc	gtc	ttc	gac	acc	gag	agc	ctg	gaa	ctg	acc	gtc	ctg	gtc	acc	5963
Gly	Ser	Val	Phe	Asp	Thr	Glu	Ser	Leu	Glu	Leu	Thr	Val	Leu	Val	Thr	
	-105					-100					-95					
gat	gcc	gcc	gcg	gtc	gag	gcc	gtg	gag	gcc	acc	ggc	gcc	ggg	acc	gag	6011
Asp	Ala	Ala	Ala	Val	Glu	Ala	Val	Glu	Ala	Thr	Gly	Ala	Gly	Thr	Glu	
	-90				-85					-80					-75	
ctg	gtc	tcc	tac	ggc	atc	gac	ggt	ctc	gac	gag	atc	gtc	cag	gag	ctc	6059
Leu	Val	Ser	Tyr	Gly	Ile	Asp	Gly	Leu	Asp	Glu	Ile	Val	Gln	Glu	Leu	
				-70					-65					-60		
aac	gcc	gcc	gac	gcc	gtt	ccc	ggt	gtg	gtc	ggc	tgg	tac	ccg	gac	gtg	6107
Asn	Ala	Ala	Asp	Ala	Val	Pro	Gly	Val	Val	Gly	Trp	Tyr	Pro	Asp	Val	
			-55					-50					-45			
gcg	ggt	gac	acc	gtc	gtc	ctg	gag	gtc	ctg	gag	ggt	tcc	gga	gcc	gac	6155
Ala	Gly	Asp	Thr	Val	Val	Leu	Glu	Val	Leu	Glu	Gly	Ser	Gly	Ala	Asp	
		-40					-35					-30				
gtc	agc	ggc	ctg	ctc	gcg	gac	gcc	ggc	gtg	gac	gcc	tcg	gcc	gtc	gag	6203
Val	Ser	Gly	Leu	Leu	Ala	Asp	Ala	Gly	Val	Asp	Ala	Ser	Ala	Val	Glu	
	-25					-20					-15					
gtg	acc	acg	agc	gac	cag	ccc	gag	ctc	tac	gcc	gac	atc	atc	ggt	ggt	6251
Val	Thr	Thr	Ser	Asp	Gln	Pro	Glu	Leu	Tyr	Ala	Asp	Ile	Ile	Gly	Gly	
	-10				-5				-1	1				5		
ctg	gcc	tac	acc	atg	ggc	ggc	cgc	tgt	tcg	gtc	ggc	ttc	gcg	gcc	acc	6299
Leu	Ala	Tyr	Thr	Met	Gly	Gly	Arg	Cys	Ser	Val	Gly	Phe	Ala	Ala	Thr	
			10					15					20			
aac	gcc	gcc	ggt	cag	ccc	ggg	ttc	gtc	acc	gcc	ggt	cac	tgc	ggc	cgc	6347
Asn	Ala	Ala	Gly	Gln	Pro	Gly	Phe	Val	Thr	Ala	Gly	His	Cys	Gly	Arg	
		25					30					35				
gtg	ggc	acc	cag	gtg	acc	atc	ggc	aac	ggc	agg	ggc	gtc	ttc	gag	cag	6395
Val	Gly	Thr	Gln	Val	Thr	Ile	Gly	Asn	Gly	Arg	Gly	Val	Phe	Glu	Gln	
	40					45					50					
tcc	gtc	ttc	ccc	ggc	aac	gac	gcg	gcc	ttc	gtc	cgc	ggt	acg	tcc	aac	6443
Ser	Val	Phe	Pro	Gly	Asn	Asp	Ala	Ala	Phe	Val	Arg	Gly	Thr	Ser	Asn	
					60					65					70	
ttc	acg	ctg	acc	aac	ctg	gtc	agc	cgc	tac	aac	acc	ggc	ggg	tac	gcc	6491
Phe	Thr	Leu	Thr	Asn	Leu	Val	Ser	Arg	Tyr	Asn	Thr	Gly	Gly	Tyr	Ala	
				75					80					85		
acg	gtc	gcc	ggt	cac	aac	cag	gcc	ccc	atc	ggc	tcc	tcc	gtc	tgc	cgc	6539
Thr	Val	Ala	Gly	His	Asn	Gln	Ala	Pro	Ile	Gly	Ser	Ser	Val	Cys	Arg	
			90					95					100			
tcc	ggc	tcc	acc	acc	ggt	tgg	cac	tgc	ggc	acc	atc	cag	gcc	cgc	ggc	6587
Ser	Gly	Ser	Thr	Thr	Gly	Trp	His	Cys	Gly	Thr	Ile	Gln	Ala	Arg	Gly	
		105					110					115				
cag	tcg	gtg	agc	tac	ccc	gag	ggc	acc	gtc	acc	aac	atg	acc	cgg	acc	6635
Gln	Ser	Val	Ser	Tyr	Pro	Glu	Gly	Thr	Val	Thr	Asn	Met	Thr	Arg	Thr	
	120					125					130					

10508.204-WO.ST25

acc gtg tgc gcc gag ccc ggc gac tcc ggc ggc tcc tac atc tcc ggc 6683  
 Thr Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly  
 135 140 145 150

acc cag gcc cag ggc gtg acc tcc ggc ggc tcc ggc aac tgc cgc acc 6731  
 Thr Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr  
 155 160 165

ggc ggg acc acc ttc tac cag gag gtc acc ccc atg gtg aac tcc tgg 6779  
 Gly Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp  
 170 175 180

ggc gtc cgt ctc cgg acc taatcgcacg ttcaatccgc tccataatcg 6827  
 Gly Val Arg Leu Arg Thr  
 185

gtcgacgcgg cggttcgcgt ccggacagca catcaccgaa atattatgga agaaaatatc 6887

agcaccatga cggccaaacg gatgcttcca acggtgctaa ctatatcacg atgtcctaca 6947

actattatca cgatcatgat aaaagctcca ttttcggatc aagtgcacagc aaaacctccg 7007

atgacggcaa attaaaaatt acgctgcacg ataaccgcta taaaaatatt gtccagcgcg 7067

cgccgagagt ccgcttcggg caagtgcacg tatacaacaa ctattatgaa ggaagcacia 7127

gctcttcaag ttatcctttt agctatgcat ggggaatcgg aaagtcacatc aaaatctatg 7187

cccaaaacaa tgtcattgac gtaccgggac tgtcagctgc taaaacgacg agcgtattca 7247

gcgggggaac ggctttatat gactccggca cggttgctgaa cggcacacag atcaacgcat 7307

cggctgcaaa cgggctgagc tcttctgtcg gctggacgcc gtctctgcat ggatcgattg 7367

atgcttctgc taatgtgaaa tcaaatgtta taaatcaagc ggggtgcgggt aaattaaatt 7427

aagaaagtga aaaacacaaa ggggtgctaac ctttgtgttt tttaattaat taaaatgttt 7487

attaacttag ttaaggagta gaatggaaaa ggggatcgga aaacaagtat ataggaggag 7547

acctatztat ggcttcagaa aaagacgcag gaaaacagtc agcagtaaag cttgttccat 7607

tgcttattac tgtcgctgtg ggactaatca tctggtttat tcccgctccg tccggacttg 7667

aacctaaagc ttggcatttg tttgcgattt ttgtcgcaac aattatcggc tttatctcca 7727

agcccttgcc aatgggtgca attgcaattt ttgcattggc ggttactgca ctaactggaa 7787

cactatcaat tgaggataca ttaagcggat tcgggaataa gaccatttgg cttatcgta 7847

tcgcattctt tatttcccgg ggatttatca aaaccggtct cgggtgcgaga atttcgtatg 7907

tattcgttca gaaattcgga aaaaaaacc tttgactttc ttattcactg ctattcagtg 7967

atttaatact ttcacctgct attccaagta atacggcgcg tgcaggaggc attatatttc 8027

ctattatcag atcattatcc gaaacattcg gatcaagccc ggcaaatgga acagagagaa 8087

aaatcggtgc attcttatta aaaaccggtt ttcaggggaa tctgatcaca tctgctatgt 8147

tcctgacagc gatggcggcg aaccgctga ttgccaagct ggcccatgat gtcgcagggg 8207

tggacttaac atggacaagc tgggcaattg ccgcgattgt accgggactt gtaagcttaa 8267

tcatcacgcc gcttgtgatt taaaactgt atccgccgga aatcaaagaa acaccggatg 8327

cggcgaaaat cgcaacagaa aaactgaaag aaatgggacc gttcaaaaaa tcggagcttt 8387



10508.204-WO.ST25

```

ccatggttat cgtgtttctt ttggtgcttg tgctgtggat ttttggcggc agcttcaaca 8447
tcgacgctac cacaaccgca ttgatcggtt tggccgttct cttattatca caagttctga 8507
cttgggatga tatcaagaaa gaacagggcg cttgggatac gctcacttgg tttgcggcgc 8567
ttgtcatgct cgccaacttc ttgaatgaat taggcatggt gtcttggttc agtaatgcca 8627
tgaaatcatc cgtatcaggg ttctcttgga ttgtggcatt catcatttta attgttgtgt 8687
attattactc tcactatttc tttgcaagtg cgacagccca catcagtgcg atgtattcag 8747
catttttggc tgtcgtcgtg gcagcggggc caccgccgct tttagcagcg ctgagcctcg 8807
cgttcatcag caacctgttc gggccaacga ctactacgg ttctggagcg gctccggtct 8867
tcttcggagc aggctacatc ccgcaaggca aatggtggtc catcggattt atcctgtcga 8927
ttgttcatat catcgtatgg cttgtgatcg gcggattatg gtggaaagta ctaggaatat 8987
ggtagaaaga aaaaggcaga cgcggtctgc ctttttttat tttcactcct tcgtaagaaa 9047
atggattttg aaaaatgaga aaattccctg tgaaaaatgg tatgatctag gtagaaagga 9107
cggctggtgc tgtggtgaaa aagcggttcc atttttccct gcaaacaaaa ataatggggc 9167
tgattgcggc tctgctggtc tttgtcattg gtgtgctgac cattacgtta gccgttcagc 9227
atacacaggg agaacggaga caggcagagc agctggcggg tcaaacggcg agaaccattt 9287
cctatatgcc gccggttaaa gagctcattg agagaaaaga cggacatgcg gctcagacgc 9347
aagaggtcat tgaacaaatg aaagaacaga ctggtgcggt tgccatttat gttttgaacg 9407
aaaaaggaga cattcgcagc gcctctggaa aaagcggatt aaagaaactg gagcgcagca 9467
gagaaatttt gtttggcggg tcgcatgttt ctgaaacaaa agcggatgga cgaagagtga 9527
tcagagggag cgcgccgatt ataaaagaac agaagggata cagccaagtg atcggcagcg 9587
tgtctgttga ttttctgcaa acggagacag agcaaagcat caaaaagcat ttgagaaatt 9647
tgagtgtgat tgctgtgctt gtactgctgc tcggatttat tggcgccgcc gtgctggcga 9707
aaagcatcag aaaggatacg ctcgggcttg aaccgcatga gatcgcggct ctatatcgtg 9767
agaggaacgc aatgcttttc gcgattcgag aagggttat tgccaccaat cgtgaaggcg 9827
tcgtcaccat gatgaacgta tcggcggccg agatgctgaa gctgcccag cctgtgatcc 9887
atcttcctat agatgacgtc atgccgggag cagggtgat gtctgtgctt gaaaaaggag 9947
aaatgctgcc gaaccaggaa gtaagcgtca acgatcaagt gtttattatc aatacgaaag 10007
tgatgaatca aggcgggag gcgtatggga ttgtcgtcag cttcaggag aaaacagagc 10067
tgaagaagct gatcgacaca ttgacagagg ttcgcaaata ttcagaggat ctcagggcgc 10127
agactcatga attttcaaat aagctttatg cgattttagg gctgc 10172

```

<210> 14  
 <211> 380  
 <212> PRT  
 <213> Artificial

<220>  
 <223> Synthetic Construct

10508.204-WO.ST25

&lt;400&gt; 14

Met Lys Lys Pro Leu Gly Lys Ile Val Ala Ser Thr Ala Leu Leu  
-190 -185 -180Ile Ser Val Ala Phe Ser Ser Ser Ile Ala Ser Ala Ala Thr Gly  
-175 -170 -165Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val Ser Met  
-160 -155 -150Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu Ala  
-145 -140 -135Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu  
-130 -125 -120Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe  
-115 -110 -105Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ala Ala  
-100 -95 -90Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu Val Ser Tyr  
-85 -80 -75Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn Ala Ala Asp  
-70 -65 -60 -55Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr  
-50 -45 -40Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Ser Gly Leu  
-35 -30 -25Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Thr Ser  
-20 -15 -10Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu Ala Tyr Thr  
-5 -1 1 5 10Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn Ala Ala Gly  
15 20 25Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val Gly Thr Gln  
30 35 40Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser Val Phe Pro  
45 50 55

Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe Thr Leu Thr

10508.204-WO.ST25  
70

60

65

Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr Val Ala Gly  
75 80 85 90

His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser Gly Ser Thr  
95 100 105

Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln Ser Val Ser  
110 115 120

Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr Val Cys Ala  
125 130 135

Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr Gln Ala Gln  
140 145 150

Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly Gly Thr Thr  
155 160 165 170

Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly Val Arg Leu  
175 180 185

Arg Thr

<210> 15  
<211> 35  
<212> DNA  
<213> Artificial

<220>  
<223> Primer

<400> 15  
ggagctctga aaaaaaggag aggataaaga atgaa

35

<210> 16  
<211> 29  
<212> DNA  
<213> Artificial

<220>  
<223> Primer

<400> 16  
gcgttccgat aatcgcggtg acaatgccg

29

<210> 17  
<211> 29  
<212> DNA  
<213> Artificial

<220>  
<223> Primer

<400> 17

10508.204-WO.ST25

ttcatgagtc tgcgccctga gatcctctg

29

<210>	18
<211>	30
<212>	DNA
<213>	Artificial

<220>  
<223> Primer

<400> 18  
taatcgcattg ttcaatccgc tccataatcg

30

<210>	19
<211>	44
<212>	DNA
<213>	Artificial

<220>  
<223> Primer

```
<400> 19
cccaacggtt tcttcattct ttatcctctc ctttttttca gagg
```

44

<210>	20
<211>	1164
<212>	DNA
<213>	Artificial

<220>  
<223> Protease 22

**<220>**  
**<221>** CDS  
**<222>** (1) .. (1164)

```
<220>
<221> sig_peptide
<222> (1)..(81)
```

```
<220>
<221> mat_peptide
<222> (577)..(1164)
```

<400> 20 45  
atg aaa aaa ccg ctg gga aaa att gtc gca agc aca gca ctt ctt  
Met Lys Lys Pro Leu Gly Lys Ile Val Ala Ser Thr Ala Leu Leu  
-190 -185 -180

att tca gtg gca ttt agc tca tct att gca tca gca gct aca gga 90  
Ile Ser Val Ala Phe Ser Ser Ser Ile Ala Ser Ala Ala Thr Gly  
-175 -170 -165

gca tta ccg cag tct ccg aca ccg gaa gca gat gca gtc tca atg 135  
Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val Ser Met  
-160 -155 -150

caa gaa gca ctg caa aga gat ctt gat ctt aca tca gca gaa gca 180  
Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu Ala  
-145 -140 -135

gaa gaa ctt ctt gct gca caa gat aca gca ttt gaa gtg gat gaa 225  
Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu  
-130 -125 -120

## 10508.204-WO.ST25

gca gcg gca gaa gca gca gga gat gca tat ggc ggc tca gtt ttt Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe -115 -110 -105	270
gat aca gaa tca ctt gaa ctt aca gtt ctt gtt aca gat gca gca gca Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ala Ala -100 -95 -90	318
gtt gaa gca gtt gaa gca aca gga gca gga aca gta ctt gtt tca tat Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Val Leu Val Ser Tyr -85 -80 -75	366
gga att gat ggc ctt gat gaa att gtt caa gaa ctg aat gca gct gat Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn Ala Ala Asp -70 -65 -60 -55	414
gct gtt ccg ggc gtt gtt ggc tgg tat ccg gat gtt gct gga gat aca Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr -50 -45 -40	462
gtt gtc ctt gaa gtt ctt gaa gga tca ggc gca gat gtt tca ggc ctg Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Ser Gly Leu -35 -30 -25	510
ctg gca gac gca gga gtc gat gca tca gca gtt gaa gtt aca aca tca Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Thr Ser -20 -15 -10	558
gat caa ccg gaa ctt tat gca gat att att ggc ggc ctg gca tat tat Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu Ala Tyr Tyr -5 -1 1 5 10	606
atg ggc ggc aga tgc agc gtt ggc ttt gca gca aca aat gca tca ggc Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn Ala Ser Gly 15 20 25	654
caa ccg ggc ttt gtt aca gca ggc cat tgc ggc aca gtt ggc aca cca Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr Val Gly Thr Pro 30 35 40	702
gtt tca att ggc aat ggc aaa ggc gtt ttt gaa cga agc att ttt ccg Val Ser Ile Gly Asn Gly Lys Gly Val Phe Glu Arg Ser Ile Phe Pro 45 50 55	750
ggc aat gat tca gca ttt gtt aga ggc aca tca aat ttt aca ctt aca Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser Asn Phe Thr Leu Thr 60 65 70	798
aat ctg gtt tca aga tat aat tca ggc ggc tat gca aca gtt gca ggc Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Ala Thr Val Ala Gly 75 80 85 90	846
cat aat caa gca ccg att ggc tca gca gtt tgc aga tca ggc tca aca His Asn Gln Ala Pro Ile Gly Ser Ala Val Cys Arg Ser Gly Ser Thr 95 100 105	894
aca ggc tgg cat tgc ggc aca att caa gca aga aat caa aca gtt agg Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn Gln Thr Val Arg 110 115 120	942
tat ccg caa ggc aca gtt tat agt ctg aca aga aca aca gtt tgt gca Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr Thr Val Cys Ala 125 130 135	990
gaa ccg ggc gat tca ggc ggc tca tat att agc ggc act caa gca caa Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr Gln Ala Gln 140 145 150	1038

## 10508.204-WO.ST25

ggc gtt aca tca ggc ggc tca ggc aat tgc agt gct ggc ggc aca aca	1086
Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Ala Gly Gly Thr Thr	
155 160 165 170	
tat tac caa gaa gtt aat ccg atg ctt agt tca tgg ggc ctt aca ctt	1134
Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser Trp Gly Leu Thr Leu	
175 180 185	
aga aca caa tcg cat gtt caa tcc gct cca	1164
Arg Thr Gln Ser His Val Gln Ser Ala Pro	
190 195	

<210> 21  
 <211> 388  
 <212> PRT  
 <213> Artificial

<220>  
 <223> Synthetic Construct

<400> 21

Met Lys Lys Pro Leu Gly Lys Ile Val Ala Ser Thr Ala Leu Leu	
-190 -185 -180	
Ile Ser Val Ala Phe Ser Ser Ser Ile Ala Ser Ala Ala Thr Gly	
-175 -170 -165	
Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val Ser Met	
-160 -155 -150	
Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu Ala	
-145 -140 -135	
Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu	
-130 -125 -120	
Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe	
-115 -110 -105	
Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ala Ala	
-100 -95 -90	
Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Val Leu Val Ser Tyr	
-85 -80 -75	
Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn Ala Ala Asp	
-70 -65 -60 -55	
Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr	
-50 -45 -40	
Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Ser Gly Leu	
-35 -30 -25	

10508.204-WO.ST25

Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Thr Ser  
           -20                  -15                  -10

Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu Ala Tyr Tyr  
      -5                  -1  1                  5                  10

Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn Ala Ser Gly  
                   15                  20                  25

Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr Val Gly Thr Pro  
                   30                  35                  40

Val Ser Ile Gly Asn Gly Lys Gly Val Phe Glu Arg Ser Ile Phe Pro  
           45                  50                  55

Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser Asn Phe Thr Leu Thr  
       60                  65                  70

Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Ala Thr Val Ala Gly  
   75                  80                  85                  90

His Asn Gln Ala Pro Ile Gly Ser Ala Val Cys Arg Ser Gly Ser Thr  
                   95                  100                  105

Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn Gln Thr Val Arg  
           110                  115                  120

Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr Thr Val Cys Ala  
           125                  130                  135

Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr Gln Ala Gln  
   140                  145                  150

Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Ala Gly Gly Thr Thr  
  155                  160                  165                  170

Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser Trp Gly Leu Thr Leu  
           175                  180                  185

Arg Thr Gln Ser His Val Gln Ser Ala Pro  
           190                  195